

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:04:29 / Search time 16.0187 Seconds

(without alignments)  
2160.495 Million cell updates/sec

Title: US-09-807-933b-5

Perfect score: 1956  
Sequence: 1 MKPFLIASAIALAVGTEM.....TYKQVCPKATKSGCSRK 360

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	702	35.9	229	2 JC7308	cellulase (EC 3.2.
2	524.5	25.8	511	2 S10527	endoglucanase B pr
3	490	25.1	393	2 S59499	cellulase eg11 - s
4	226	11.6	471	1 A26160	cellulose 1,4-beta
5	226	11.6	471	1 A38979	cellulose 1,4-beta
6	198.5	10.1	475	2 S49886	probable membrane
7	196	10.0	662	2 A45155	mucin P1M-C.1 - Af
8	179.5	9.2	888	2 T46726	secreted acid phos
9	173	8.8	328	2 S67570	hypothetical prote
10	173	8.8	1217	2 S52714	sericinB - silkw
11	168.5	8.6	371	2 S20075	promastigote surfa
12	168	8.6	1832	2 T31113	mucin-like glycopr
13	164	8.4	797	1 VGBEX1	glycoprotein X pre
14	164	8.4	2761	2 T21064	hypothetical prote
15	163.5	8.4	402	2 E86185	hypothetical prote
16	163.5	8.4	605	2 T33913	hypothetical prote
17	162.5	8.3	418	1 S28372	cellulase (EC 3.2.
18	162.5	8.3	1367	2 S51959	hypothetical prote
19	162	8.3	796	2 T21460	hypothetical prote
20	161	8.2	477	2 S53362	mucin 5AC (clone J
21	160.5	8.2	786	1 A47547	serine proteinase
22	160	8.2	327	2 S20074	promastigote surfa
23	160	8.2	503	2 S63257	probable membrane
24	160	8.2	1275	2 T33369	hypothetical prote
25	158	8.1	354	2 T46740	microfilament shea
26	158	8.1	1161	2 S57180	probable membrane
27	157.5	8.0	344	2 T40167	hypothetical prote
28	156.5	8.0	263	2 S01360	salivary glue prot
29	156.5	8.0	750	2 T42614	probable envelope

30	156	8.0	438	1 S70602	cellulose 1,4-beta
31	155.5	7.9	786	2 T16509	hypothetical prote
32	155.5	7.9	1251	2 T21389	hypothetical prote
33	155	7.9	217	2 S01358	salivary glue prot
34	154.5	7.9	1032	2 T34433	hypothetical prote
35	154.5	7.9	2232	2 T34434	hypothetical prote
36	153.5	7.8	725	2 A41258	a-egg-luciferin core
37	153	7.8	299	2 S56031	pathogenesis-relat
38	153	7.8	1360	2 T33932	hypothetical prote
39	152.5	7.8	798	2 T34248	hypothetical prote
40	151.5	7.7	971	2 B90835	probable tail fibre
41	151.5	7.7	973	2 C85693	probable membrane
42	151	7.7	1566	2 T20058	hypothetical prote
43	150	7.7	770	2 T22808	hypothetical prote
44	149	7.6	605	2 S48940	hypothetical prote
45	148	7.6	446	2 T45525	Wsc4 homolog (limp

## ALIGNMENTS

## RESULT 1

UC7308 cellulase (EC 3.2.1.4) - Scopulariopsis brevicaulis

N:Alternate names: endoglucanase I

C:Species: Scopulariopsis brevicaulis

C>Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000

C:Accession: UC7308; PC7087

R:Nakatani, F.; Kawaguchi, T.; Takada, G.; Sumitani, J.; Moriyama, Y.; Arai, M.

Biosci. Biotechnol. Biochem. 64, 1238-1246, 2000

A:Title: Cloning and sequencing of an endoglucanase gene from Scopulariopsis brevicaulis

A:Reference number: UC7308

A:Accession: UC7308

A:Molecule type: DNA

A:Residues: 1-229 <NAK>

A:Experimental source: strain TOF-1212

A:Accession: PC7087

A:Molecule type: protein

A:Residues: 21-377,149-164 <NA2>

A:Gene: eg1

A:Introns: 147/3

C:Keywords: glycosidase; hydrolase

Query Match	Score	Match	Length	DB ID	Description
152	ASGNETTRVYDCCPKSCSWPGKADVT-SPVGSCKNDGKTLADNN-TONGC-VGGSSYTC	35.9%	229	2	JC7308
209	NDNQPWVSDDLAYGFAAASISGSEATWCCACEELFTSTAVYKXQVNTGSDLG	35.9%	229	2	S10527
21	ASGCTTTRVYDCCPKSCSWPGKADVT-SPVGSCKNDGKTLADNN-TONGC-VGGSSYTC	35.9%	229	2	S59499
81	SSHPWAVDDELAYGMAVNIIGQTESDWCACVLEFTTGAVGSKMIVQATNTGSDLG	35.9%	229	1	A26160
269	SNTGAHPLQMPGGGCVGYNGCATQWGAFTMGWGRVGVSSASDLSLPSALQAGCKWR	35.9%	229	1	A38979
141	NN--HFDIAMPGGGVGFNCTDQWSPPMGMDRQVHTRADCSFPEALKAGCEWR	35.9%	229	2	S49886
329	PGWPNADNPMTVYKQVTCPRATKSGCSR	35.9%	229	2	A45155
198	FDWFGTDNPVSVFVEVCPALVQSGCOR	35.9%	229	2	T46726

## RESULT 2

S10527 endoglucanase B precursor - Pseudomonas fluorescens

C:Species: Pseudomonas fluorescens

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999

C:Accession: S10527

R:Gilbert, H.J.; Hall, J.; Hazlewood, G.P.; Ferreira, L.M.A.

Mol. Microbiol. 4, 759-767, 1990

A>Title: The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspect  
A:Reference number: S10527; MUID:90355836; PMID:2117693  
A:Accession: S10527  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-511 <GIL>  
A:Cross-references: EMBL:X52615; NID:g45497; PID:CAA36844.1; PID:g45498  
C:Superfamily: bacterial cellulose-binding domain homology; glycosidase GWM domain hom  
F:31-128/Domain: bacterial cellulose-binding domain homology <BCB>  
F:180-217/Domain: glycosidase GWM domain homology <GWM>  
F:32-127/Disulfide bonds: #status predicted

Query Match 26.8%; Score 524.5; DB 2; Length 511;  
Best Local Similarity 30.8%; Pred. No. 2.6e-25;  
Matches 123; Conservative 59; Mismatches 119; Indels 99; Gaps 12;

QY 33 CGGKMDGPTCCSGSTCYDYPNPFYSCVNPENLSTNSKHTTTEASKTTTG 92  
DB 127 CGGGSSABSVASSSS-----SSSVSSSTPRSSSSSVSPGTSSSS 171  
QY 93 SKTTTTEASK-----KT-----TTEASKTT 115  
DB 172 SSVLTGAQCMNVGTLTPLCNNTNGWYEDRSVCARTTGAQAPYGVIVTSSTPLS 231  
QY 116 TTBASKTTTTHKASTTSSSSASTYSAVSGASNGETTRWDCCKPSCMPGKA 175  
DB 232 SSSSSSVASSSSSSATSSASVS-SVPPIDGCGC-NGVATRYWDCCKHCGMSANV 288  
QY 176 -VTSFVSGCNKDGKTLANNQNGCVGSSYTCNDNQPWVSDLAFFAASISGSE 234  
DB 289 PSLVPSLPGCSANNRLSDVSVSGSCDGGGYCMWKIPFVSPILATGYAITS-SGDV- 346  
QY 235 ATWCAACFELTFT-----STAVKGMVVOVNTGSDLGSGTGAHFDLQMPGG 283  
DB 347 ---CGRCYQLOFTGSSYNAPGDPGSAALAKTMIVATIGVY---SGGQDILVPGG 400  
QY 284 VGTTCATONGAPTDGMAKRGG-----VSASDC-----SNLPS 319  
DB 401 VAFAPNCSHMVGVSNAELGAQYGFPLAACKOOLGYNASISQYKSVCLNRCDSVFGRLT 460  
QY 320 ALQAGCKRFPKPNADNPTMYKQVTCPKAITAKSGCSR 359  
DB 461 QLOGCTWPAEMFEADNPDLKKEVPCPAELITTSGMNR 500

RESULT 3  
S59499  
cellulase egII - smut fungus (Ustilago maydis)  
C:Species: Ustilago maydis (corn smut)  
C>Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 28-May-1999  
C:Accession: S59499  
R:Schauwecker, F.; Wanner, G.; Kahmann, R.  
Biol. Chem. Hoppe-Seyler 376, 617-625, 1995  
A>Title: Filament-specific expression of a cellulase gene in the dimorphic fungus *Ustilago*  
A:Reference number: S59499; MUID:96145728; PMID:8590631  
A:Accession: S59499  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-393 <SCH>  
A:Cross-references: GB:S81598; NID:g1478378; PID:AA836147.1; PID:g1478379

Query Match 25.1%; Score 490; DB 2; Length 393;  
Best Local Similarity 46.1%; Pred. No. 2.7e-23;  
Matches 101; Conservative 29; Mismatches 67; Indels 22; Gaps 9;

QY 156 GETTYMDCKRSCGMPGKADVTSPVSGCNKDKTLADN---NQNQNGVSSSTYCDN 211  
DB 27 GHAIRYWDCLLASASWEGAPVAPADKADGVTLLDSKDPSSGSGCNKCKMCSGM 86  
QY 212 QPWWVSD--LAVGFAPASISGSEATWCCACFELTF---TSTVKGKMWVOVNTG 265  
DB 87 QPFDDETPTLAFGCA--FTTGQESDPTCACFYAEFEHDAQGMKMKRLIFQVTTNG 144

QY 266 DLGSNTGAHFDLQMPGGVGIY-NGCATONGAPTDGMAKRGVSSASDCSNLPSALQAG 324  
DB 145 DVQSQ---NFPQRIQGGGLGAFPKGCPAQMOWEASLWMDQYGVKASATGCKLPRLQEG 201  
QY 325 CKMRFG-WFKANDNPTM--TYKQVTCPEAITAKSGCSCK 360  
DB 202 CKMRSEW---GDNPVLKGSPEKRVKCPKSLIDRSGCCPK 237

RESULT 4  
A26160  
cellulose 1,4-beta-cellulosidase (EC 3.2.1.91) II precursor - fungus (*Trichoderma reesei*)  
N:Alternate names: exo-cellulohydrolase II  
C:Species: *Trichoderma reesei*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A26160; NID:274-278, 1987  
R:Chen, C.M.; Gitzali, M.; Stafford, D.W.  
Bio/Technology 5, 274-278, 1987  
A>Title: Nucleotide sequence and deduced primary structure of cellulohydrolase II from  
A:Reference number: A26160  
A:Accession: A26160  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-471 <CH>  
R:Teeri, T.T.; Lehto-Vaara, P.; Kauppinen, S.; Salovuori, I.; Knowles, J.  
Gene 51, 43-52, 1987  
A>Title: Homologous domains in *Trichoderma reesei* cellulytic enzymes: gene sequence an  
A:Reference number: A26472; MUID:87248061; PMID:3596237  
A:Accession: A26472  
A:Molecule type: DNA  
A:Residues: 1-471 <TE>  
A:Cross-references: GB:M16190; NID:g170540; PID:AAA4210.1; PID:g170541  
C:Genetics:  
A:Introns: 31/2; 160/2; 243/1  
C:Superfamily: cellulose 1,4-beta-cellulosidase II; fungal cellulose-binding domain hom  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
F:31-62/Domain: fungal cellulose-binding domain homology <FCB>

Query Match 11.6%; Score 226; DB 1; Length 471;  
Best Local Similarity 25.5%; Pred. No. 5.9e-07;  
Matches 94; Conservative 41; Mismatches 156; Indels 78; Gaps 13;

QY 5 TIASAIIALAVGTEMAHAECSKAYVQCGKMDGPTCCSGSTCYDYPNPFYSCVP 64  
DB 8 TLATLATLAASVPLERQA---CSSVWGCGGQNMSPGCCASGTCV--YENDYSGCLP 63  
QY 65 NENLSTNSKSHKTTTESAKTTTGSKTTTTEASKTTTTEASKTTTTEASKTT 124  
DB 64 G---AASSSTRASTTRVSPPTSSRS--SATPPGSTTRVPVVGSGTATYSGNPFV 118  
QY 125 TTYKASTSTSSSSSAS---TVNSAVSGASGGETTRWDCCKPSCMPGKADVTSP 180  
DB 119 GVTTPANAYVASEVSLAIPSLTGAMTAFAAVK-----VPSFMMLDLDTKPL 168  
QY 181 VGSCKDKGTLADNNTQNGCVGSS---YTCNDNQPWVSDLAFFAASISGSEA 235  
DB 169 M-----EQLTADRINKNGNVAQGFVVYDLPDRCCALASNGEY-----SIADGVA 217  
QY 236 TW-----CCACFELTFTSTAVKGMVVOVNTGSDLGSGTGAHFDLQMPGGGCI 286  
DB 218 KYKAYIDTIRQIVVEYSIRFLVIEPDSLNLVNTLGPICANAGSY-----DEC 269  
QY 287 YNGCATONGAPTDGMAKRGVSSASDCSNLPSALQAGCKMRFPKANDNPTM---TY 342  
DB 270 INVAVTQNLN-----NVAMVLIDHGHGMLGMPANQDPAQLFANVY 311  
QY 343 KQVTCPEAI 351  
DB 312 KNASPRAL 320

RESULT 5  
A38979

Query Match	11.6%;	Score 226;	DB 1;	Length 471;
Best Local Similarity	25.5%;	Pred. No. 5.9e-07;		
Matches	94;	Conservative 41;	Mismatches 156;	Indels 78; Gaps 13

```

OY      5 TTASSAILALAVGEMAHABCSKAYVOCGGKMMGPCCESGTCVDYEPDNPFYSCVP 64
Db      8 TLATLATLAAAPLEEROA--CSSVWGCGGGMWGSPCCASGSTCV--YSNDYISQCLP 63
OY      65 NENILSTWKS1HKTTTESAKKTTTGKSKTTTTEASKTTTTEASKTTTTEASKTTT 124
Db      64 G---AASSSSSTRASTTSRVSPTTSRS--SATPPGSTTTTRVPVSGITVYSGNPFV 118
OY      125 TTTTKASTSTSSSSSSAS---TNSAVSGASGNGETTRWDCCKPSCSWPGKADVTSF 180
Db      119 GVTPEANNAVYASEVSLSAIPSLTGAMAT2AAAVAK-----VPSMTMDITD3KIPL 168
OY      181 VGSCKDKGKTLADNNTONGCVGSSS----YTCNDNQPWVVSZDLAYGPAASISGSSBA 235
Db      169 M-----EOTLADITANKNGNVYGO4FVYVLPBPDCAALASNGEY-----SIADGVA 217
OY      236 TW-----CCACFELTPTSTAVKGGKMMVQVNTGSDLSNTGAHE5DLOMPGGVGVI 286
Db      218 KYKNYIDTIRQIVVEYSIDIRTLVIEPDSLANLVNLTGPKANNQSAV-----LEC 269
OY      287 YNGCATONGAPPTDGGARYGVGSASDCSNLPSALQAGCKMRF6GMFKNADNPTM---TY 342
Db      270 INVA7TQNLNLP-----NVAMYLIDAGHAGW8GM9PANDP10PAALQ11FANVY 311
OY      343 KQVTCPKAI 351
Db      312 KNASSPRL 320

```

```

RESULT 6
S49886
probable membrane protein YIL123w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein Y18277.06
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C:Accession: S49886
R:Hamlyn, N.; Churcher, C.
submitted to the EMBL Data Library, November 1994
A:Reference number: S49881
A:Accession: S49886
A:Molecule type: DNA
A:Residues: 1-475 <HAN>
A:Cross-references: GB:Z47047; EMBL:Z46833; NID:G603997; PID:G763223; MIPS:YIL123w
C:Genetics:
A:Gene: SGD:SIM1
A:Cross-references: SGD:S0001385; MIPS:YIL123w
A:Map position: 9L
C:Superfamily: Saccharomyces NCA3 protein
C:Keywords: transmembrane protein
F:82-98/Domain: transmembrane #status predicted <TM>

```

```

OY      1 MKFLT-----IASAIIA-----LAVGTEHMAHAAACSKY---YQCGKNMGPPCTCEGSS 48
Db      1 MKFLTAVTTLLISSAIVSALPHVDVHODEDAHQHKRAVAKYVTEVVVDSJGHTTPPAAS 60
OY      49 -----TCVDYP-----DNPEYOCVPIENLTS 70
Db      61 EVATPAATSAITTTSVLAPTSAAAGIAASIVSSAALAKMEKIDBAASATASTSGASS 120
OY      71 TNKSHKTTTTEASAKTTTYTGSKKTTTTEASKKTTTTEASK--KTTTTEASKKTTTYYTK 128
Db      121 SSSSSSMTSTLESSVSSSEEAAPTSTVTSATQTSASATKSSSTSTSTSTSTST 180
OY      129 KASTSTSSSSSSASTNTNVAVSGASG-----NGETRYWDCCKBSCSPGKADVTP--V 181
Db      181 TSTSTSSSSSSSSSSSGSGSIYGLDLADPSGPEKQDGTIPECKRPPSGQVISIDWI 240
OY      182 GSCMKDKITLADNNTQNGCVGSS--STYCNUNOP-----WVSDDLAYGFAASI----- 229
Db      241 GEGMGSGEINTDSTGTGSCKEGSCYSC--OQMSKTOW--PSPQPDGDSVGLLCKN 296
OY      230 -----SGSEATWCC-----ACFELTFTSTAYK-----GKXWVYQ--TWTGSDL----- 267
Db      297 GYLIRSNTDADYLTCEWGEVAAYVWSKUSKGAICRTDPTGETENNVIPTVEBGSSLLPLTV 356
OY      268 -----GSNTGAHFLOMPGCGVGIVYNGCATOWG-----APTDGMCARGAYGV 308
Db      357 VDQDPTTYTWEKKTSAQY--YVNNAGAVSVEBGCI--WQTSGSGIGMAPLNPGASTSGV 412
OY      309 S-----SASDCSNLPSALOAGCKMFRGMFK--NADNPTMYTKOYT 346
Db      413 TYLSLIPNPNNSDALNPNVKIIVAADSSNVIGB-----CVYENGESGAGDCTVS--VT 465
OY      347 CPKA 350
Db      466 SGKA 469

```

```

RESULT 7
A4515
mucin FTM-C.1 - African clawed frog (fragment)
C|Species: Xenopus laevis (African clawed frog)
C|Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000
C|Accession: A4515
R.Hausser, P.; Hoffmann, W.
J. Biol. Chem. 267, 24620-24624, 1992
A>Title: P-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FTM-C.1)
A|Reference number: A4515; MUID:93077556; PMID:1447205
A|Accession: A4515
A|Status: preliminary
A|Molecule type: mRNA
A|Residues: 1-662 <HNU>
A|Cross-references: GB:02115; NID:g214147; PIDN:AAA74725.1; PID:g951460
C|Superfamily: trefoil homology
F:162-202/Domain: trefoil homology <TRF1>
F:307-347/Domain: trefoil homology <TRF2>
F:354-394/Domain: trefoil homology <TRF3>
F:526-566/Domain: trefoil homology <TRF4>
F:573-613/Domain: trefoil homology <TRF5>
F:621-661/Domain: trefoil homology <TRF6>

Query Match      10.0%; Score 196; DB 2; Length 662;
Beet Local Similarity 26.1%; Pred. No. 5.8e-05;
Matches 89; Conservative 33; Mismatches 127; Indels 92; Gaps 15;

Qy      32 QCGGKNDGPYCSGSST-----CVDPDNPYSQCVPENILTSNKSKTTTESAKK 86
       |||||          |||         |::|   :|||:|
Db      373 QCGRKN-----CCPDSSISGTWC-----PSTSQAATPKTTTTPTTTTTPTTTTK 420
       |||||          |||         |::|   :|||:|

87 TTTTSGSKTTTTEASKTTTTEASKTTTTEASKTTTTEASKTTTTEASKTTTTEASKTTT 145

```





